Large visual neuron assemblies receptive fields estimation using a super-resolution approach

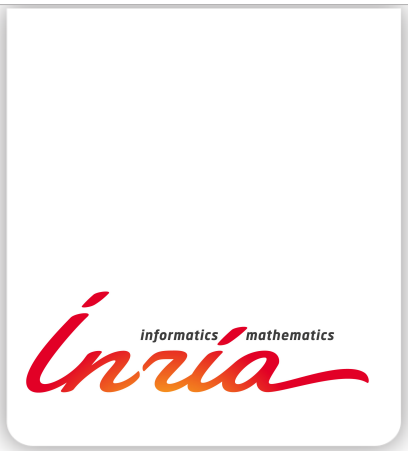
Daniela Pamplona, Gerrit Hilgen, Matthias H. Hennig, Bruno Cessac, Evelyne Sernagor, Pierre Kornprobst

Large visual neuron assemblies receptive fields estimation using a super-resolution approach

Daniela Pamplona, Gerrit Hilgen, Matthias H. Hennig, Bruno Cessac, Evelyne Sernagor, Pierre Kornprobst

RESEARCH REPORT
N° 9383
December 2020
Project-Team Biovision

ISSN 0249-6399 ISRN INRIA/RR--9383--FR+ENG



Large visual neuron assemblies receptive fields estimation using a super-resolution approach

Daniela Pamplona*†, Gerrit Hilgen‡, Matthias H. Hennig§,
Bruno Cessac*, Evelyne Sernagor‡, Pierre Kornprobst*

Project-Team Biovision

Research Report n° 9383 — version 2 — initial version December 2020 —
revised version February 2021 — 21 pages

This work was partially supported by the EC IP project FP7-ICT-2011-9 no. 600847 (RENVISION)

* Université Côte d'Azur, Inria, France
† Ecole Nationale Supérieure de Techniques Avancées, Institut Polytechnique de Paris, U2IS, 828 Boulevard des Marchaux, 91120 Palaiseau, France
‡ Institute of Neuroscience, Faculty of Medical Sciences, Framlington Place, Newcastle upon Tyne, NE2 4HH, United Kingdom
§ Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh, EH8 9AB, United Kingdom

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Abstract: Computing the spike triggered average (STA) is a simple method to estimate the linear receptive fields (RFs) of sensory neurons. For random, uncorrelated stimuli the STA provides an unbiased RF estimate, but in practise white noise is not a feasible stimulus as it usually evokes only very weak responses. Therefore, for a visual stimulus, it is often used images of randomly modulated blocks of pixels. This solution naturally limits the resolution at which an RF can be obtained. Here we show that this limitation can be overcome by using a simple super resolution technique. We define a novel type of stimulus, the Shifted White Noise (SWN), by introducing random spatial shifts in the usual stimulus in order to increase the resolution of the measurements. In simulated data we show that the average error using the SWN was 1.7 times smaller than when using the classical stimulus, with successful mapping of 2.3 times more neurons, covering a broader range of RF sizes. Moreover, successful RF mapping can be achieved with short recordings of about one minute of activity, more than 10 times more efficient compared to the classical white noise stimulus. In recordings from mouse retinal ganglion cells with large scale microelectrode arrays, we could map 18 times more RFs covering a broader range of sizes. In summary, here we show that randomly shifting a the usual white noise stimulus significantly improves Rf estimation, and requires only short recordings. It is straight forward to extend this method into the time dimension and adapt it to other sensory modalities.

Key-words: Receptive fields, sensory neurons, Spike Triggered Average, super-resolution, retina

Estimation des champs récepteurs de grands ensembles de neurones visuels par une approche de super-résolution

Résumé :

La méthode du Spike Triggered Average (STA) est une méthode simple pour estimer la partie linéaire des champs récepteurs (CR) des neurones sensoriels. Avec des stimuli aléatoires et non corrélés, la STA fournit une estimation non biaisée des CR, mais en pratique, le bruit blanc n'est pas un stimulus efficace car il ne provoque généralement que des réponses très faibles. C'est pourquoi, pour un stimulus visuel, on utilise souvent des images de blocs de pixels modulés de façon aléatoire. Cette solution limite naturellement la résolution à laquelle un CR peut être obtenue. Nous montrons ici que cette limitation peut être surmontée en utilisant une simple technique de super résolution. Nous définissons un nouveau type de stimulus, le bruit blanc décalé (SWN), en introduisant des décalages spatiaux aléatoires dans le stimulus habituel afin d'augmenter la résolution des mesures. Dans les données simulées, nous montrons que l'erreur moyenne en utilisant le SWN était 1,7 fois plus petite que lors de l'utilisation du stimulus classique, avec une cartographie réussie de 2,3 fois plus de neurones, couvrant une plus large gamme de tailles de CR. De plus, une cartographie CR réussie peut être obtenue avec des enregistrements courts d'environ une minute d'activité, plus de 10 fois plus rapidement qu'avec le stimulus classique de bruit blanc. Dans les enregistrements de cellules ganglionnaires rétiniennes de souris obtenues avec des réseaux de microélectrodes à grande échelle, nous avons pu cartographier 18 fois plus de CR couvrant une plus large gamme de tailles. En résumé, nous montrons ici que le décalage aléatoire d'un stimulus de bruit blanc habituel améliore considérablement l'estimation des CR et ne nécessite que de courts enregistrements. Il est facile d'étendre cette méthode à la dimension temporelle et de l'adapter à d'autres modalités sensorielles.

Mots-clés : Champs récepteurs, neurones sensoriels, Spike Triggered Average, super-résolution, rétine

Introduction

Sensory neurons are characterised by their Receptive Field (RF), which is the area of the sensory space they respond to upon stimulation. In visual neurons it is the area of the visual field these cells respond to when light intensity changes. Estimating the size and shape of an RF with high accuracy requires measurements at sufficiently high spatial resolution. Ideally, RF measurements should consist of sampling at very high resolution, which means stimulating small subunits (pixels) of the RF in sequence. However, when these pixels are too small, neural response are less likely as normally cells respond to simultaneous stimulation of many pixels in their RF. On the other hand, when the pixel size is too large, responses do not reflect RF sizes faithfully. This problem is exacerbated by the fact that RF sizes are not homogeneous across the neuronal population. For example, many Retinal Ganglion Cells (RGCs) types have smaller RFs in the centre of the retina than in the periphery, hence the optimal pixel size to determine central RFs is smaller than for measurements in the periphery. Owing to new technological developments in recording approaches consisting of large-scale, high-density multielectrode arrays (MEAs), it is now possible to record responses to light from hundreds to thousands of neurons simultaneously [2], encompassing both central and peripheral cells. Such an experimental scenario requires designing new stimuli that can yield high-resolution measurements for all cells across the neural population.

In this study, we present a novel approach to measure RFs at high fidelity from large and heterogeneous neural populations recorded simultaneously. The classical way of estimating RFs is to estimate the Spike Triggered Average (STA) from evoked neural recordings. In short, STA estimates the average stimulus before a spike. If the stimulus is white noise and the recording sufficiently long, then this average corresponds to the neuron's RF [20, 9]. The white noise stimulus consists of a series of non-overlapping binary images shown successively in time, with individual images showing a black or white pixel of similar size presented in random order but with equal probability. This stimulus, here termed Basic White Noise (BWN), has the size of the blocks as parameter.

In the case of single-cell recordings, the block size is defined according to the experimenter's expectations: it must be smaller than the expected RF size in order to yield high-resolution measurements, but it cannot be too small, in order to avoid weak neural responses. Various more or less heuristic approaches could, in principle, fulfill these conditions. For instance, one can start from a tiny block size and gradually increase it during the experiment in the direction of the larger stimulus-neural response correlation [8] or mutual information [14, 15]. However, in the case of large neuronal populations, one cannot merely apply the same procedure as for individual cells. Indeed, a optimal block size for one neuron will be sub-optimal for another one in the population. Similarly, a high stimulus-neural response correlation (or mutual information) for one neuron might be low for another one. As a consequence, experimenters must design stimuli that best suit the population as a whole regarding to the degree of heterogeneity in the neuronal population under investigation. Here we use a novel white noise stimulus, which we call Shifted White Noise (SWN). The size of each block is large, ensuring strong responses from all RGCs. However, the blocks are shifted only by a fraction of their size, yielding high-resolution sampling. In other words, large pixels ensure strong light responses while sub-pixel shifts yield super-resolution measurements, enabling us to measure all RFs with great accuracy.

Super-resolution is a class of image processing methods to estimate a high-resolution image from a set of low-resolution ones [28]. Super-resolution has been successfully applied to a variety of domains such as, e.g., video [13], remote sensing [27, 17], and medical imagery [10, 22]. Interestingly, even though the present study does not address pure image reconstruction issues, we found interesting analogies. In a nutshell, to increase the resolution of RFs, the idea is not to

diminish the block size (which would end up with weaker neural activity, and thus, faulty RF estimation), but instead, keep a high block size while introducing additional spatial variability in the stimulus to improve the accuracy of the responses. Our method ensures ensures that the majority of cells rapidly respond to the stimulus (large block size), while at the same time, the resulting RFs have much higher resolution (block shift).

A novel class of visual stimuli

The classical stimulus to estimate RFs, here called BWN, is illustrated in Fig. 1 (A). It consists of a sequence of binary images showing equal-sized blocks, with colors drawn randomly from a Bernoulli distribution with a probability 0.5. Each image is displayed for a fixed time. Since the STA relies on averaging stimuli within time windows, the resulting spatial precision (resolution) of the estimated RF is equal to the BWN's block size. In order to increase resolution of the measurements, the most obvious approach would be to decrease the block size, as shown in Fig. 1 (B). However, this decreases the response of most neurons and computing the STA reliably would require much longer recordings.

The super-resolution approach used here preserves large block sizes to guarantee stronger responses, but to randomly shift the binary images in space at each presentation. Two such examples are shown in Fig. 1(C–D). In isolation these shifted blocks will not change responses compared to BWN blocks of similar size. However, combining these low-resolution estimates with the shifts will yield significantly high-resolution RFs.

This allows using the following strategy. Let us denote by β the block size of the stimuli, and define the target resolution for the RF as $\alpha = \beta/k$, where k denotes the increase in resolution (e.g. double resolution: $k = 2$). The target resolution defines a baseline shift of the same value from which one can define a series of random spatial shifts $s = n\alpha$ with $n \in \{0, \ldots, k-1\}$. Considering blocks of BWN images, one shift is applied to each block, yielding a succession of blocks characterized by the same shift size, but in random direction. Using the STA for each block provides one RF per block. These are then combined to get a high-resolution RF.

Instead of applying the STA for each block, we randomize the different shifts and apply the STA globally. The sequence of images obtained in this way is what we call SWN. It is illustrated in Fig. 1(E). The main advantage of SWN has over BWN is that the high-resolution RF directly results from the STA. There is no need to estimate intermediary low-resolution RFs and then combine them.

In summary, while the resolution of the RF is given by the block size with BWN, it is independent of block size with SWN. Instead, it is provided by the baseline shift. This provides the experimenter with the option of choosing a sufficiently large block size to increase activity levels in response to the stimulus. Since the resolution is given by the block shift size, the only remaining limitations may be technical such as the size of the projected pixels, something inherent to the experimental set up and light stimulation equipment. Another significant advantage of SWN is that it introduces more variability, favoring better light responses across the overall population. In other words, we expect not only to obtain RFs of high-resolution, but we also expect to be able to define the RFs for more cells over shorter stimulation periods.

From now on, in this paper, the stimuli are named as follows: BWN-Bβ denotes a BWN with a block size β; SWN-Bβ-Sα denotes a SWN with a block size β and a baseline shift α. All sizes are expressed in μm.

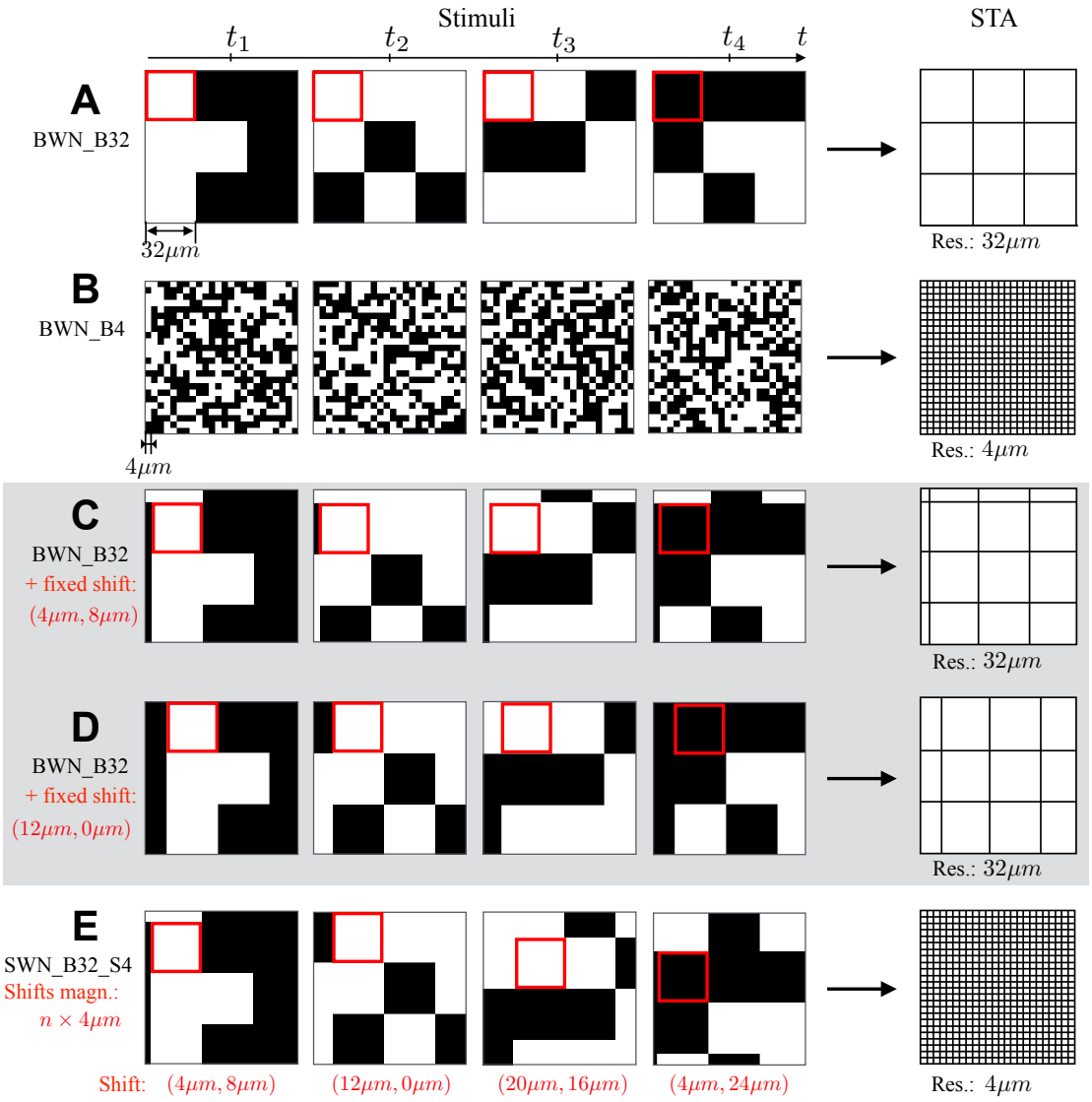


Figure 1: Exploring variation across white noise stimuli: From BWN to SWN. (A) BWN-B32: BWN with blocks of 32 µm . In the BWN case, the resolution of the RF is the same as the size of the blocks. (B) BWN-B4: same as (A) but with blocks of size 4 µm. (C) and (D): Series of BWN-B32-like stimuli based on example (A), with an additional fixed spatial shift. The spatial shift is represented by a red square showing how the upper-left block has been moved. The resolution of the STA is still the same as the size of the blocks. However, the different shifts will infer different samplings of the RF that could be combined to achieve a high-resolution RF (see text for details). Note that starting from (A) was chosen only for explanatory purposes so that readers could compare both conditions. The binary patterns should be a priori random. (E) SWN-B32-S4: SWN with blocks of 32 µm and random spatial shifts using a baseline shift of 4 µm. With this condition, it is the baseline shift that defines the resolution of the RF.

Methods

Synthetic data: Stimuli

In the experimental analysis we consider three stimuli: two BWNs of low and high resolution (32 µm and 4 µm, respectively) and one SWNs whose STA is of resolution 4 µm (with blocks of 160 µm and shifts of 4 µm).

For each stimulus we generated $27,000$ images of 88×88 pixels, where one pixel corresponds to a square of size $16\,\mu m^2$. Here we arbitrarily set the origin at the central pixel of the spatial domain, which will be more convenient to specify neurons' population. Images were refreshed at 30.3Hz, meaning that, each image was presented for 33 ms.

Each stimulus was fed independently to our artificial population of neurons described below to obtain the simulated spiking output and then the RF estimation using STA. Each image was presented for 1ms corresponding to one discrete time-step in our model.

Synthetic data: Artificial retinal Ganglion cell model

In this work, we define a population of neurons described by Linear-Nonlinear Poisson (LNP) models [19, 6]. These functional models are widely used by experimentalists to characterize the cells that they record, map their RFs, and characterize their spatio-temporal feature selectivities [12, 7, 4, 25, 1]. The STA of a LNP neuron stimulated with random white noise converges to the RF of the neuron, up to a multiplicative constant [20].

In its simplest form, a LNP model is a convolution of the stimulus L with a spatio-temporal kernel K followed by a static non-linearity and stochastic (Poisson-like) mechanisms of spikes generation. Here we use this model to simulate RGCs' spiking activity in response to our stimuli. We consider that K is a Difference-of-Gaussians centered at (c_x, c_y), center size σ_c and surround size σ_s. The detailed definition is given in the Appendix.

Synthetic data: Neural population construction

Given the neuron's model described in above, our goal is to define an heterogeneous population of such neurons that will cover the different possible experimental scenarios but, for the sake of simplicity, we do not consider orientation or direction selective cells. Heterogeneity comes from the choice of the parameters in the spatial part of kernel K (see (4)), namely the center of the RF, (c_x, c_y), and the size of the central Gaussian, σ_c.

The generation process is illustrated in Fig. 2. First, we choose as a reference grid the grid provided by the BWNs of low resolution, i.e., with blocks of size 32 µm. The idea is to define a population of neurons where we vary positions and kernel sizes. More precisely, we define a population of 216 neurons as follows:

- We first define a set of positions to evaluate the consequences of the alignment of the RF with the block center. Starting from position $c_0 = (0, 0)$ (at the center of the spatial domain), we define a family of neurons equally sampled along the diagonal direction in steps of one pixel ($\delta_c = 4\,\mu m$) until the next block center c_{max} (see Fig. 2(A)). Doing so, we define nine positions:

$$(c_x, c_y) = \{(0,0), (\delta_c, \delta_c), \ldots, (8\delta_c, 8\delta_c)\} = \{(0,0), (4,4), \ldots, (32,32)\}.$$

  Note that one parameter is sufficient to describe neuron position, namely we use c_x in the remaining of this paper.

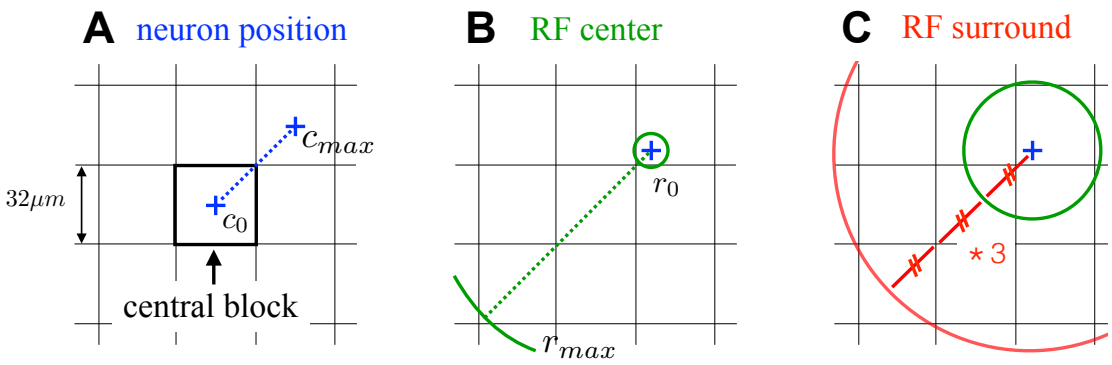


Figure 2: Stages to build the artificial RGC population. (A) We define nine positions equally distributed between the center of the central block (c_0) and the center of the adjacent block in the diagonal direction (c_M). (B) For the spatial kernel, we define 24 possible values of σ_c, so that the positive part radius vary from half pixel to 1.5 blocks in steps of half pixel. The smallest RF have radius of 2 µm and the largest 24 µm. (C) We assume here that the surround variance is three times higher than center variance.

- Then, for each neuron position, we define a family of neurons with varying spatial kernel sizes (see Eq. (4)). The smallest radius r_0 corresponds to a center standard deviation $\delta_\sigma = 0.784$ µm. In this way, the size of the center Gaussian (the circumference radius where the Difference of Gaussian (DOG) changes from positive to negative) is 2 µm. The remaining σ_c were defined as a multiple of this one, increasing in steps of 2 µm (meaning, half pixel) the center radius (see Fig. 2(B)). Doing so, we defined 24 possible values of σ_c:

$$\sigma_c = \{\delta_\sigma, 2\delta_\sigma, \ldots, 24\delta_\sigma\} = \{0.784, 1.568, \ldots, 18.824\}.$$

  Concerning the surround standard deviation, σ_s, it was set to three times σ_c as usually fixed in the literature (see Fig. 2(C)). Thus only one parameter is sufficient to describe the spatial kernel amplitude, i.e., σ_c.

Synthetic data: STA

The STA is a reverse-correlation technique commonly used to estimate the RF of neurons that relies on both neuron model and stimulus properties [4, 25, 20, 9]. The STA corresponds to the average sequence of images preceding spikes. It is defined as follows. Consider a neuron (model or experimental) that spiked at times $t_1, t_2, \ldots, t_n$ when stimulated by a spatio-temporal stimulus $S(x, y, t)$, then the STA of this neuron, denoted by $A(x, y, \tau)$, is given by:

$$A(x, y, \tau) = \frac{1}{n} \sum_{i=1}^{n} S(x, y, t_i - \tau) \tag{1}$$

with x and y belonging to the same spatial domain as the stimulus S and τ is on $\{-\mathcal{T} \ldots 0\}$ where $-\mathcal{T}$ defines the temporal support of the STA.

STA allows a parameter free estimation of the RF, it is easy to design and use experimentally in any biological sensory modality. To estimate the STA, we use the PRANAS, an open and free platform for retinal analysis and simulation [3].

Synthetic data: Evaluation

We propose three criteria to assess the validity of the SWN.

Criterion 1 (Number of mapped RFs) We consider that a RF was mapped if the STA 2D spatial profile was structured, thus if there is a blob, or a point where non zeros values are concentrated. To recess automatically whether a RF was mapped or not, we made a Student's T-test with a significance level $\alpha = 1e^{-8}$ on the 2D-spatial profile of each STA. If the 2D-spatial profile is Gaussian distributed (meaning, if it is noise), then RF was not mapped. Otherwise, we considered that the RF was mapped.

Criterion 2 (RF parametric description) For a mapped RF, we fit the spatial RF with a DOG as defined in eqn. (4). For that, we used the Trust Region Reflective method [18, 16], which is a bounded minimization algorithm. In practice, in order to balance between parameters variability and algorithm efficiency, we defined large bounds for each parameter: σ_c lower value is 0.1 and higher value is three times the image size in µm; c_x and c_y lower value is equal to 1 µm and higher value the image length. To avoid local minima each STA was fitted 12 times with different initializations uniformly sampled within the bounds. Then, for the analysis, we selected the parameters that minimize the fitting error.

Criterion 3 (STA error) In the synthetic case, we can compare the STAs with the GTs "point-by-point". To do so, we measure the angle between the two vectors using the cosine similarity as suggested in [20]. If we denote by A the output of the STA, this angle is given by:

$$E(\bar{K}, A) = cos^{-1} \frac{< \bar{K}, A >}{\|\bar{K}\| \ \|A\|\|}, \tag{2}$$

where $\bar{K}$ is the the true neuron's kernel defined by (3), and $\|.\|$ the usual euclidean norm.

Experimental data: Stimuli

We have used four stimuli to map RGC RFs in mouse retinal wholemounts: two BWNs of low and high resolution (160 µm and 40 µm, respectively) and two SWNs whose STA is of high and super high resolution (for both stimuli blocks of 160 µm and shifts of 40 µm and 4 µm, respectively).

We generated 60,000 images of each stimulus grouped in 20 blocks of 3000. These blocks were randomly sorted before projecting onto the retina to avoid response bleaching bias for certain stimulus conditions. Each image was 664×664 pixels, where one pixel corresponds to a square of size $4\,\mu m^2$. Images were refreshed at 30.3Hz (showed for 33 ms).

Light stimuli were projected onto the retina as described previously in [23] and attenuated using neutral density filters to high mesopic light levels (mean luminance 11 cd/m2).

Experimental data: MEA recordings

All experimental procedures were approved by the ethics committee at Newcastle University and carried out in accordance with the guidelines of the UK Home Office, under control of the Animals (Scientific Procedures) Act 1986. Experiments were performed as described previously in [11].

Briefly, a female mouse (aged 42 days) was dark-adapted overnight and killed by cervical dislocation. Eyes were enucleated, and following removal of the cornea, lens, and vitreous body, they were placed in artificial cerebrospinal fluid (aCSF) containing the following (in mM): 118

NaCl, 25 $NaHCO_3$, 1 NaH_2PO_4, 3 KCl, 1 $MgCl_2$, 2 $CaCl_2$, 10 glucose, and 0.5 l-Glutamine, equilibrated with 95% O_2 and 5% CO_2. The retina was isolated from the eye cup and flattened for MEA recordings. All procedures were performed in dim red light, and the room was maintained in darkness throughout the experiment. Retinal recordings were performed on the BioCam4096 platform with BioChips 4096S+ (3Brain GmbH, Lanquart, Switzerland), integrating 4096 square microelectrodes (21×21µm, pitch 42 µm) on an active area of 2.67×2.67mm.

The spatial extent ($7.12\,mm^2$) of the MEA chip allowed us to record simultaneously from large retinal areas (see Fig. 6). The small electrode pitch (42 µm) enables sampling from many individual RGCs from these areas, providing us with an unbiased very large analytical sample size. After recording, spikes were sorted and the rasters were formed. Single-unit spikes were sorted using the T-Distribution Expectation-Maximisation algorithm in Offline Sorter (Plexon Inc, Dallas, USA).

Experimental data: STA

Before computing the STAs, the stimulus images (independently of the stimulus) were cropped at (640×640) pixels to remove the partial blocks and resized to the smallest size possible without compromising the results. Precisely, the images of BWN-B160 were reduced to (16×16) pixels, BWN-B40 and SWN-B160-S40 were reduced to (40×40) pixels, while SWN-B160-S4 was not resized. This allowed to reduce the population STA computation time from days to several minutes (in BWN-B160) or hours (BWN-B40 and SWN-B160-S40). The STA estimation was performed as for the synthetic data.

Experimental data: Evaluation

All responsive cells were considered for evaluation. The evaluation was performed similarly to the synthetic data. First the STA was estimated and analysed using Criterion 1 (Number of mapped RFs). For the neurons whose RF was successfully mapped, Criterion 2 (RF parametric description) was considered. However, here only the center Gaussian was considered, as is usually done when analyzing mouse retinal RFs [11]. Criterion 3 could not be used since Ground Truth (GT) was not available.

Results

Synthetic data: Single neuron level

Here we focus on only one neuron chosen arbitrarily from the population. This neuron is located in $(c_x, c_y) = (4\delta_c, 4\delta_c)$ which is at the intersection of blocks, and with a central variance kernel $\sigma_c = 24\delta_\sigma$. Note that this was chosen to have the most favorable situation for BWN-B32 (as shown later in Fig. 4).

For each of the three stimulus types, we used ten instances of each stimulus to generate rasters, with a total of 30 rasters. Each instance of the stimulus represents 20,000 images, which generated rasters of 11 minutes long.

Neural responses for each stimuli are indicated in Tab. 1. As expected, stimuli with larger blocks (BWN-B32 and SWN-B32-S4) induced stronger responses. In addition, note that SWN-B32-S4 generates the strongest response. This fact is relevant because, as shown in [21], the STA error decreases as a function of the number of spikes.

In Fig. 3(A) we show the spatio-temporal kernels estimated with each stimuli. These results can be compared qualitatively with the ground truth, here represented with a spatial discretization of $4\mu m$). With BWN-B32, the result lacks precision in space. It has strong vertical and

Stimulus	# spikes	spike rate
BWN-B32	9108	13.8Hz
BWN-B4	6204	9.4Hz
SWN-B32-S4	9438	14.3Hz

Table 1: Average neuron response to each stimuli across trials, with the number of spikes and the corresponding spike rate. Note that, for all the stimuli, the spike rate is high because we do not considered a refractory period in the neuron model. The spike rate using SWN-B32-S4 is 1.5 times higher than using BWN-B4

horizontal edges corresponding to the stimulus block size. The temporal part, however, is properly estimated. With BWN-B4, when the block size is smaller, the results are very noisy in both space and time, and no relevant information can be detected. With SWN-B32-S4, the spatial aspect has the same resolution as the ground truth, and this method gives goods result in terms of precision too, with good estimate of the RF shape, even if some weak, noisy patterns remain in the periphery. The temporal part is accurate as well, as with BWN-B32. These observations suggest that SWN allows increasing the spatial part of the kernel's spatial resolution with the same stimulation time.

In Fig. 3(B) we show how the error (eqn. 2) evolves in time for each condition (Criterion 3). Estimates are done at every minute of stimulation time, i.e., at time t, STA is estimated based upon spikes recorded in the time window $[0, t]$. The SWN error is always the smallest followed by the BWN-B32 and then by BWN-B4. Note that, even when using the BWN-B32 for an extremely long time, the STA error will not converge to 0, due to the different resolutions between the STA and the ground truth.

Finally, in Fig. 3(C), we estimate the number of times that the neuron's RF was mapped by the end of each minute of stimulation. With BWN-B4, the RF was never mapped (before 11 minutes). In contrast, the number of times that the RF was mapped with BWN-B32 increases with the stimulation time to reach 100% at seven minutes. With SWN-B32-S4, this number rapidly grows to reach 100% at one minute already. In other words, successful RF mapping of single neurons was achieved 10 times faster with the SWN than with BWN.

Taken together, these observations on a single neuron suggest that our approach allows better quality RF estimates and that these estimates need less stimulation time. In the next section, we consider an entire population of neurons to verify whether these observations hold in more generally.

Synthetic data: Population level

For the population of 216 neurons defined in the methods Section, the three stimuli were presented for 11 minutes. Note that since inter-trial variability is low (see Fig. 3(B), we use only one trial in this section.

In Fig. 4(A), we show the error between STA and the ground truth (Criterion 3), after 11 minutes over the whole population. This compact representation gives very instructive insights into the precision reached by each stimulus and how such performance depends on the neurons' characteristics. First, globally, the average error over the entire population is 55.6, 83.6, and 48.7 degrees for BWN-B32, BWN-B4, and SWN-B32-S4, respectively. In other words, for the same resolution, 4 µm, the average error using SWN is 1.7 times smaller than using BWN. Then, when delving into more details, starting with BWN-B4, we observe that strong errors are made for most neurons in the population, due to the slow rate of convergence with this stimulus, as observed in the single neuron case. We note a minor exception for the smallest RF, which seems to be

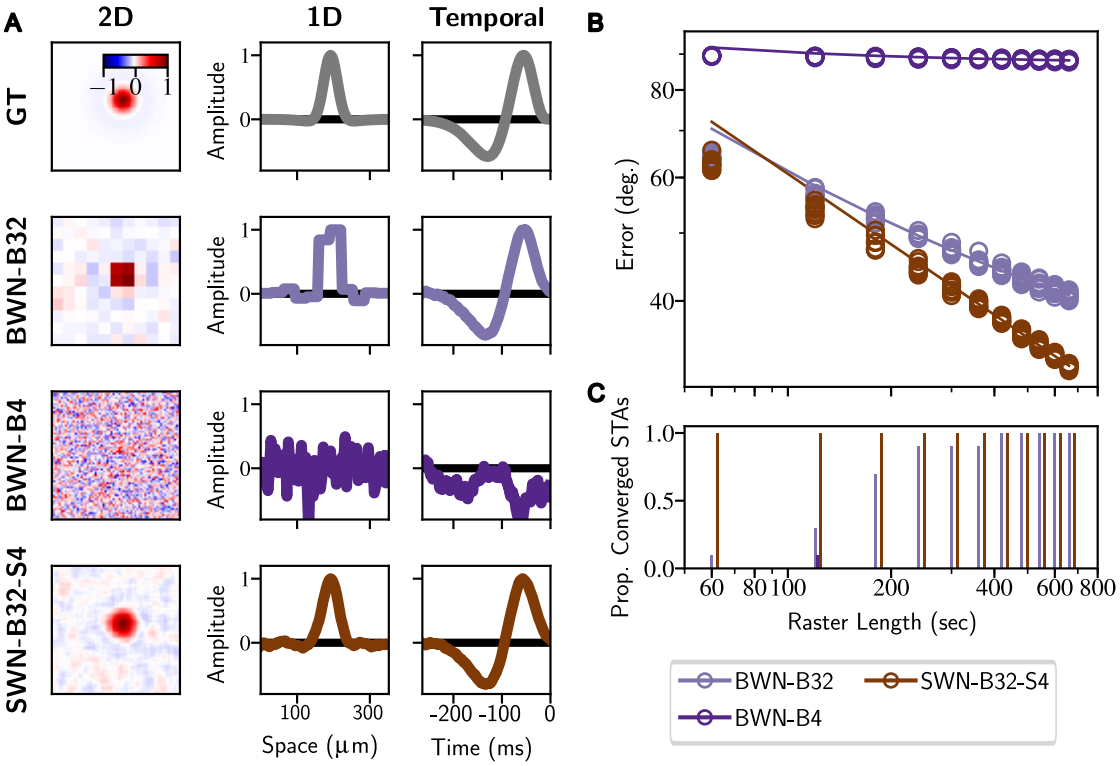


Figure 3: Single neuron analysis. (A) Comparison of spatio-temporal RF estimations. (2D) Spatial slice taken at the point with maximal amplitude of the STA. (1D) Horizontal cut from 2D slice passing through the point with maximal amplitude. (Temporal) Temporal cut from the spatial-temporal STA passing through the point with maximal amplitude. The SWN allows for increasing the spatial part quality while preserving the quality of the temporal profile (B) Error as a function of time between the RF estimate and ground truth. An estimate is done at each minute (circles, for each trial) and we fit this error with a power law (continuous line). The SWN always performs better than BWN, and the convergence rate is faster. (C) Proportion of mapped RFs at each minute. It is noticeable that very early, from the first minute, 100% of RF can be mapped with SWN. To reach 100%, one must wait seven minutes with BWN-B32. For BWN-B4, convergence is much slower so that no RF can be fit before 11 minutes.

better-captured when using small block sizes. Considering larger block size, namely BWN-B32, the error decreases as the RFs become larger, which reflects matching between block and RF sizes. Another effect emerges, namely some level of dependence on the neurons' position relative to the block. When the position of a neurons is "in-between" blocks, errors become larger. Which is presumably related to the fact that neurons in such positions experience less spatial average variations. This effect completely disappears with SWN-B32-S4. Indeed, although errors keep decreasing with increasing RF sizes, the neuron's position does not affect the accuracy of the responses anymore, which offers great advantages when dealing with experimental data.

From the 216 neurons, applying the Criterion 1 at the end of stimulation, all neurons were mapped with both BWN-B32 and SWN-B32-S4, but only 90 neurons were mapped with BWN-B4. In other words, for the same resolution, the SWN mapped 2.3 times more neurons than the BWN.

In Figs. 4(B)–(C), we compare the parameters of the fitted kernel with the original ones (Criterion 2), illustrating kernel parameters at 5 (left) and 11 minutes (right) respectively. These panels complement Fig. 4(A) in two ways: (1) they provide an interpretation of the nature of the error, and (2) they give an idea of the evolution in time of this error. Both BWN-B32 and SWN-B32-S4 result in biased fittings, towards a larger center size than the ground truth's. For BWN-B32, this bias essentially remains over time, while for SWN-B32-S4, it decreases (Fig. 4(B)). Furthermore, BWN-B32 bias depends on the RF position, while with the increase of the raster length the SWN-B32-S4 bias becomes independent of the position (Fig. 4(C)). Concerning BWN-B4, there is no bias, at least for the very small RFs that could be mapped. At the same time, no medium and large RFs could be mapped with this stimulus.

Experimental data: Single neuron level

In Fig. 5 we show four representative cases of estimated RFs. For these four neurons, we observe different situations concerning the number of mapped RFs measured with Criterion 1. In all cases selected, SWN-B160-S4 was mapped. No striking difference appears for SWN stimuli with shifts of 40 µm and 4 µm (when both were mapped). Results with BWN-B40 are always noisy, even when the RF is mapped. We also found several analogies with the syntetic data. RFs estimated with the SWN stimuli were smoother than RFs estimated with the BWN. The RF temporal profile, on the other hand, was not altered by the shifting process.

Experimental data: Neural Population

Fig. 6(A) illustrates the log spiking activity of the retina recorded form the RGC layer. Responses to light occurred across the entire active area of the MEA, with particular emphasis on the dorsal-lateral axis.

In Fig. 6(B) shows the distribution of the centers of all mapped RFs depending on the stimulus type. This was achieved by fitting STAs with DOGs to find their center position and size. Overall, the distribution of the STA's center was similar to the activity map for BWN-B160 and both SWN. However, when looking closer, RF centers of BWN-B160 tended to be less uniformly distributed, more concentrated in the ventral direction.

Fig. 7(A)–(B) show STA convergence properties with respect to the four stimuli. Basic count of the number of receptive fields found (Fig. 7(A)) reveals that the SWNs method over performs the BWNs method in all conditions. Interestingly, increasing the resolution leads to different behaviors. With BWN to increase the resolution yields to fewer cells with mapped RFs. With SWN to increase the resolution yields to in even more cells with successfully mapped RFs. When the final STA resolution is fixed to 40 µm and we use the SWN, we can successfully map 18 times more RFs than when using the BWN. Going a step further, Fig. 7(B) shows which cell were

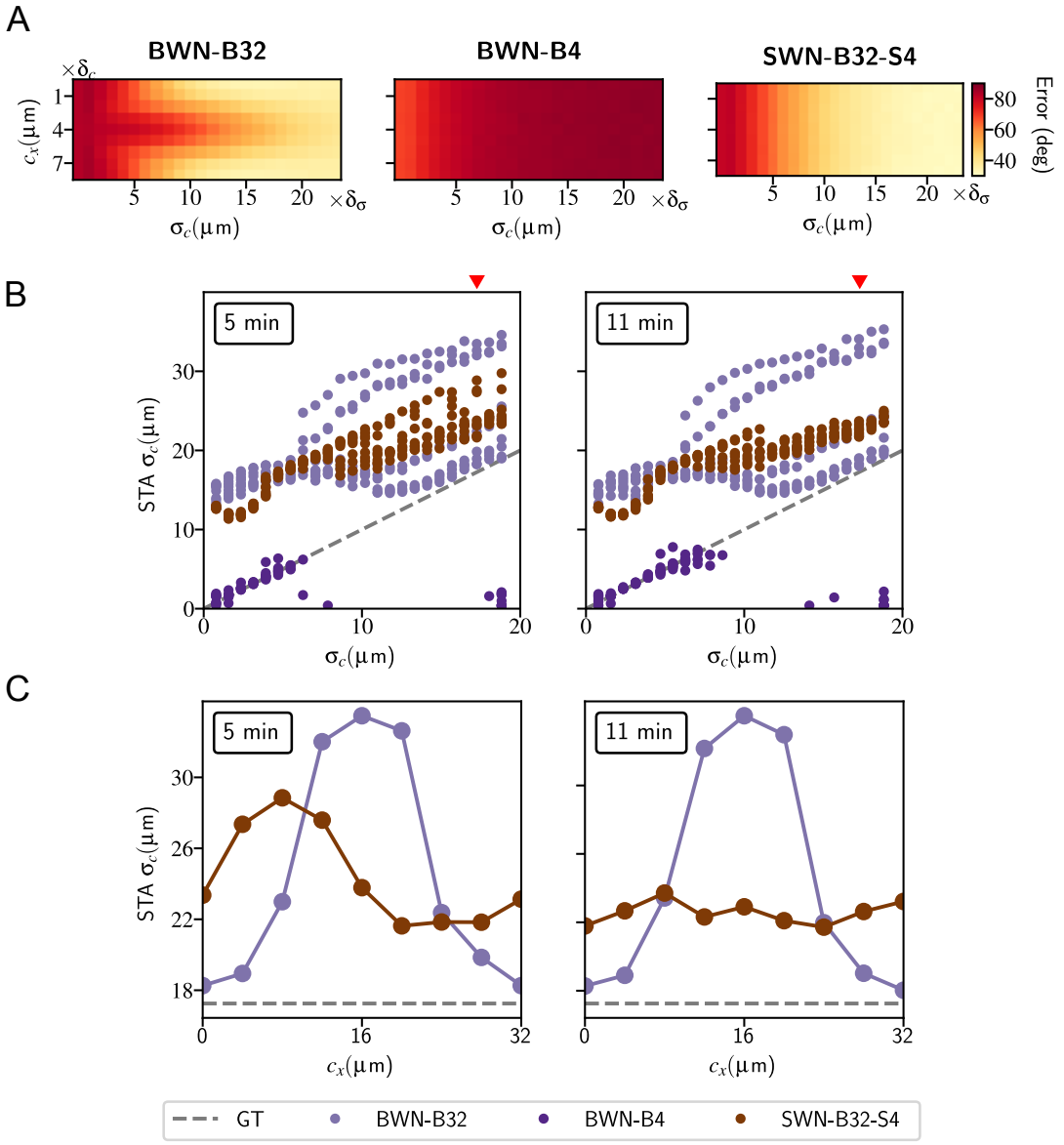


Figure 4: Population analysis. (A) Error between the RF estimate and ground truth, as a function of position and size. Results show that SWN-B32-S4 offers the best performance at the population level, with no dependence on the neurons' position. (B)–(C) Comparing the fitted kernel parameters with the original ones after five minutes and 11 minutes, respectively. (B) Estimated sizes against the ground truth sizes for all RFs. On the left at five min of stimulation, on the right at 11 min of stimulation. (C) Estimated sizes in function of the RF position for the neurons with $\sigma_c = 17.25$. On the left at five min of stimulation, on the right at 11 min of stimulation. Results show that stimuli with large block sizes give biased RF sized. These biases remain in time but tend to become more uniform with SWN-B32-S4 contrarily to BWN-B32 where they always depend on the neuron's position. With a small block size, i.e., BWN-B32, the situation is different. No bias is observed, but this is only true for a minimal number of neurons for which the RF was mapped.

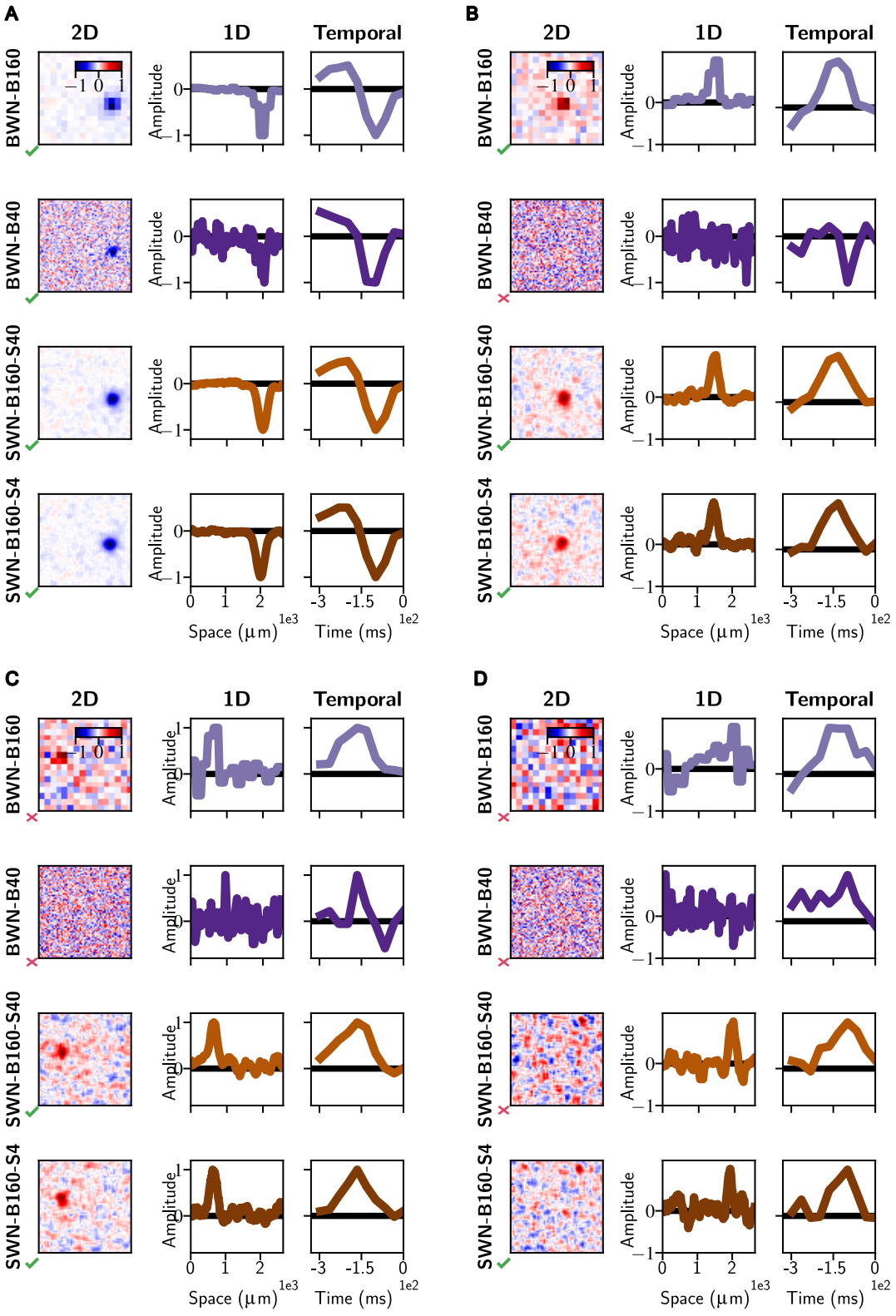


Figure 5. STAs representatives examples for four neurons, showing different situations in terms of RF mapping depending on the stimulus. The representation is the same as in Fig. 3. A green tick icon (resp. a red cross icon) is used to indicate that the RF was mapped or not. In general, SWN yields to smoother spatial STAs than the BWN, without changing the temporal STA.

Retina 9383

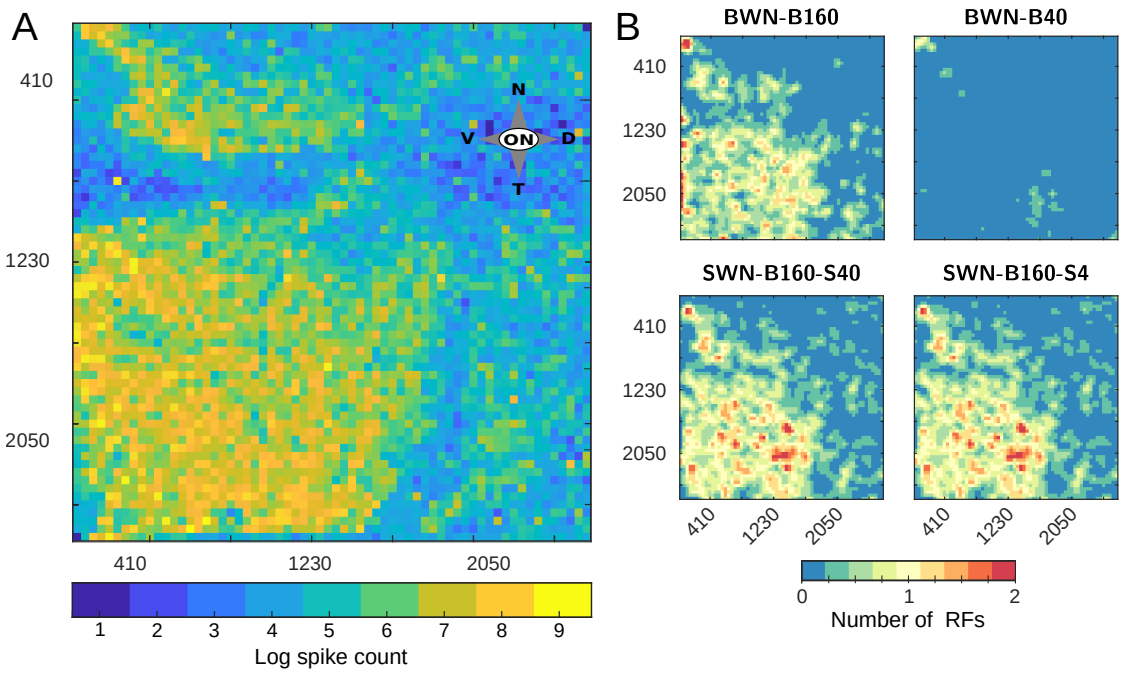


Figure 6: Retinal activity pan-retinal view. (A) The Log spike count during the entire experiment for each retinal channel. This results in a visualisation of the retina outline and gives an overall estimation of the number of active channels. (B) Number of RFs distributed over the MEA. Retina orientated identically to panel (A). For visualization purposes the values of all sub-panels were smoothed with a Gaussian kernel. RFs' positions are in agreement with the activity map.

found for each stimulus. This is represented as an Euler diagram. Except for a few cells, we observe the following pattern of inclusions:

$$\text{BWN-B40} \subset \text{BWN-B160} \subset \text{SWN-B160-S40} \subset \text{SWN-B160-S4}.$$

Note that a small percentage of RFs' was mapped with SWN-B160-S40 and not with SWN-B160-S4, but this number is four times smaller than the reverse situation.

In Fig. 7(C) we compare the estimated center sizes for the mapped RFs. RFs with radius smaller than 40 µm or larger than 180 µm were considered outliers as their size is not biologically plausible (see, for instance, [11]). Outliers are not shown in this Fig. The number of outliers depended on the resolution. While high-resolution stimuli lead to a relatively small number of outliers: 5, 7 and 6 for BWN-B40, SWN-B160-S40 and SWN-B160-S4, respectively. In contrast, using the low resolution BWN-B160 stimulus yielded to 287 outliers. As in the synthetic case, the BWN of high resolution was pruned for small centre sizes, while the BWN of low resolution and the SWN-B160-S40 stimuli were pruned to larger radii. Notably, the SWN-B160-S4 was not pruned to a specific range. Nevertheless, the shape of the radii distribution depended on the stimulus as well. The spread around the preferred value was low in the BWN-B40, medium in the BWN-B160 and SWN-B160-B40 cases and large in the BWN-B160-B4 case. These results are quantified in Tab. 2. BWN-B40 yielded the smallest RF size values (mean and standard deviation), while values are equally large when using BWN-B160 and SWN-B160-S40. Using SWN-B160-S4 yields the largest standard deviation.

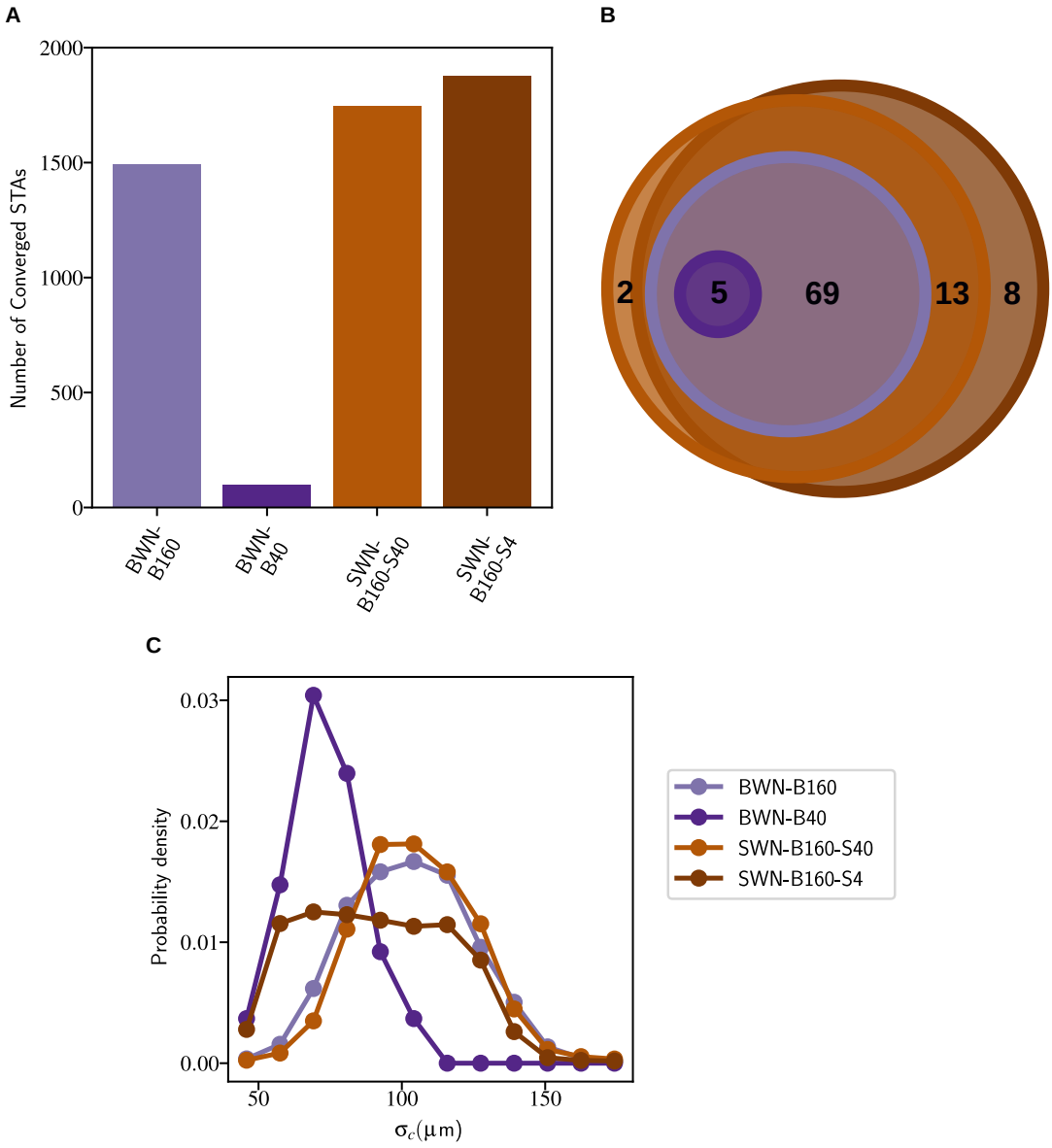


Figure 7: Mapped RFs statistics. (A) Number of mapped RFs per stimulus. (B) Euler diagram (in percentage) of mapped RFs. Cell percentages below 2% are not shown. (C) Distribution of RFs sizes per stimulus. For the same resolution, 40 µm, SWN mapped 18 times more RFs then BWN. Furthermore, SWN mapped a broader range of sizes.

stimulus	RF size
BWN-B160	102.7 ± 21.3
BWN-B40	74.2 ± 13.3
SWN-B160-S40	104.9 ± 19.9
SWN-B160-S4	91.5 ± 25.4

Table 2: Mean and standard deviation statistics of RFs sizes per stimulus

Discussion and conclusion

In summary, here we show that a shifted white noise stimulus considerably improves RF estimation for retinal ganglion cells. Using synthetic data, we demonstrate that with SWN, RF estimation is independent of the position of individual neurons relative to the stimulus, which is not the case for BWN. The resolution is always higher with SWN, since there is no compromise between responsiveness (given by the block size) and resolution, and the latter increases by reducing the stimulus baseline shift. At the population level, not only do we achieve higher mapping resolution, but we also map more RFs, with more neuronal variability.

Because neurons exhibit stronger activity when presented with larger block sizes, the use of SWN stimuli also leads to faster RF mapping, which makes the STA approach more efficient. In the case of synthetic data, we showed that the whole population can be mapped seven times faster with SWN. This is important when dealing with experimental data because mapping RFs with STA is often just one preliminary step in a much longer experimental pipeline (using various stimuli), which often leads to bleaching of light responses (see, e.g., [23])).

Another advantage of SWN is that the new stimuli are also easy to produce and the same reverse correlation methods can be used to recover the RF. Of course, in practice, one still needs to choose a suitable block size. This choice still relies on the experimenter expertise, but with SWN, it is less critical since the variability introduced by the shift will compensate for a suboptimal block size value. A possible strategy is to use a block size smaller than the expected RF field size or the dendritic field width of the measured cell types with a shift about 1/5 of the block size. Nevertheless, this strategy must be adapted to the experimental context such as animal species, cell type, recording conditions and the study goal.

Computationally, we found that the STA mapped with the SWN are larger than the ground truth, but this bias seems to decrease with stimulation time. In contrast, similar bias was found in the low resolution BWN, but without changing significantly with time.

Visually, there were no striking differences between the RFs mapped with the SWNs-B160-S40 and SWN-B160-S4. However, their fitting sizes were not similar. With SWN-B160-S4 a broader distribution of sizes were mapped than with SWN-B160-S40. SWN-B160-S4 mapped both small and large RFs. Remarkably, to map both small and large RFs was not possible with any other stimuli, thus with the one pixel resolution SWN we mapped the larger amount of RFs and without the sizes being biased towards a specific range.

The general approach applied to STA in this paper — making use of super-resolution methods to boost the performance of RF estimation methods — will allow for more efficient stimuli design in sensory physiology. For example, a similar approach could be applied in the temporal domain by randomizing each frame's presentation time in the stimulus. We also expect that this general approach could be applicable to other spike-triggered methods like the Spike Triggered Covariance [20, 26], since it is a good approximation of Gaussian White Noise, spherical and easy to implement. Furthermore, this super-resolution idea might also be useful on other sensory modalities where the STA has been shown to be interesting.

References

[1] Atencio, C.A., Sharpee, T.O., Schreiner, C.E.: Cooperative nonlinearities in auditory cortical neurons. Neuron **58**(6), 956–966 (2008)

[2] Berdondini, L., Imfeld, K., Maccione, A., Tedesco, M., Neukom, S., Koudelka-Hep, M., Martinoia, S.: Active pixel sensor array for high spatio-temporal resolution electrophysiological recordings from single cell to large scale neuronal networks. Lab on a Chip **9**(18), 2644–2651 (2009)

[3] Cessac, B., Kornprobst, P., Kraria, S., Nasser, H., Pamplona, D., Portelli, G., Viéville, T.: Pranas: A new platform for retinal analysis and simulation. Frontiers in Neuroinformatics **11**, 49 (2017). DOI 10.3389/fninf.2017.00049. URL `http://journal.frontiersin.org/article/10.3389/fninf.2017.00049`

[4] Chichilnisky, E.J.: A simple white noise analysis of neuronal light responses. Network: Comput. Neural Syst. **12**, 199–213 (2001)

[5] Croner, L.J., Kaplan, E.: Receptive fields of P and M ganglion cells across the primate retina. Vision research **35**(1), 7–24 (1995)

[6] Dayan, P., Abbott, L.F.: Theoretical neuroscience: computational and mathematical modeling of neural systems. Computational Neuroscience Series (2001)

[7] DeAngelis, G.C., Ohzawa, I., Freeman, R.D.: Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. i. general characteristics and postnatal development. Journal of neurophysiology **69**(4), 1091–1117 (1993)

[8] Földiák, P.: Stimulus optimisation in primary visual cortex. Neurocomputing **38-40**, 1217–1222 (2001)

[9] Gollisch, T.: Estimating receptive fields in the presence of spike-time jitter. Network: Computation in Neural Systems **17**(2), 103–129 (2006)

[10] Greenspan, H.: Super-resolution in medical imaging. Computer Journal **52**(1), 43–63 (2009). DOI 10.1093/comjnl/bxm075

[11] Hilgen, G., Pirmoradian, S., Pamplona, D., Kornprobst, P., Cessac, B., Hennig, M.H., Sernagor, E.: Pan-retinal characterisation of light responses from ganglion cells in the developing mouse retina. Scientific reports **7**(1), 1–14 (2017)

[12] Jones, J.P., Palmer, L.A.: The two-dimensional spatial structure of simple receptive fields in cat striate cortex. Journal of Neurophysiology **58**(6), 1187–1211 (1987)

[13] Liu, C., Sun, D.: A bayesian approach to adaptive video super resolution. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 209–216 (2011)

[14] Machens, C.K.: Adaptive sampling by information maximization. Physical review letters **88**(2), 228104 (2002)

[15] MacKay, D.J.C.: Information-based objective functions for active data selection. Neural Computation **4**, 590–604 (1992)

[16] The Mathworks, Inc., Natick, Massachusetts: MATLAB version 9.3.0.713579 (R2017b) (2017)

[17] Merino, M.T., Nunez, J.: Super-resolution of remotely sensed images with variable-pixel linear reconstruction. IEEE Transactions on Geoscience and Remote Sensing **45**(5), 1446–1457 (2007)

[18] Moré, J.J., Sorensen, D.C.: Computing a trust region step. SIAM Journal on Scientific and Statistical Computing **4**(3), 553–572 (1983)

[19] Odermatt, B., Nikolaev, A., Lagnado, L.: Encoding of luminance and contrast by linear and nonlinear synapses in the retina. Neuron **73**(4), 758–773 (2012)

[20] Paninski, L.: Convergence properties of three spike-triggered analysis techniques. Network **14**, 437–464 (2003)

[21] Paninski, L.: Maximum likelihood estimation of cascade point-process neural encoding models. Network: Comput. Neural Syst. **15**(04), 243–262 (2004)

[22] Peeters, R.R., Kornprobst, P., Nikolova, M., Sunaert, S., Vieville, T., Malandain, G., Deriche, R., Faugeras, O., Ng, M., Van Hecke, P.: The use of super-resolution techniques to reduce slice thickness in functional mri. International Journal of Imaging Systems and Technology **14**(3), 131–138 (2004)

[23] Portelli, G., Barrett, J.M., Hilgen, G., Masquelier, T., Maccione, A., Di Marco, S., Berdondini, L., Kornprobst, P., Sernagor, E.: Rank order coding: a retinal information decoding strategy revealed by large-scale multielectrode array retinal recordings. eNeuro **3**(3) (2016)

[24] Rodieck, R.W.: Quantitive analysis of cat retinal ganglion cell response to visual stimuli. Vision Research **5**(12), 583–601 (1965)

[25] Schwartz, O., Pillow, J.W., Rust, N.C., Simoncelli, E.P.: Spike-triggered neural characterization. Journal of vision **6**, 484–507 (2006)

[26] Simoncelli, E.P., Paninski, J.P., Pillow, J., Schwartz, O.: Characterization of neural responses with stochastic stimuli. The cognitive neurosciences, 3rd edition pp. 327–338 (2004)

[27] Wang, H., Wen, D.: The progress of sub-pixel imaging methods. In: J. Ojeda-Castaneda, S. Han, P. Jia, J. Fang, D. Fan, L. Qian, Y. Gu, X. Yan (eds.) Selected Papers from Conferences of the Photoelectronic Technology Committee of the Chinese Society of Astronautics: Optical Imaging, Remote Sensing, and Laser-Matter Interaction 2013, vol. 9142, pp. 166–170. International Society for Optics and Photonics, SPIE (2014)

[28] Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., Zhang, L.: Image super-resolution: The techniques, applications, and future. Signal Processing **128**, 389–408 (2016)

Supplementary material

LPN neuron model

The LNP model used in the simulation part (see Methods Section) has three stages:

Stage 1 describes how the neuron integrates stimulus intensity over space and time. The stimulus is denoted by $S(x, y, t)$ where $(x, y) \in \{0 \dots M, 0 \dots N\}$ and $t \in \{0, 1, 2, \dots\}$. The spatio-temporal kernel of the neuron is denoted by $K(x, y, t)$ where $(x, y) \in \{0 \dots M, 0 \dots N\}$ and $t \in \{0, \dots, T\}$ (i.e., T is its temporal support) The resulting integration denoted by $L(t)$ is defined by an inner product in space and a convolution in time:

$$L(t) = \sum_{\tau=0}^{T} \sum_{x=0}^{N} \sum_{y=0}^{M} K(x, y, \tau) S(x, y, t - \tau).$$

Then we assume that the kernel K is separable is space and time, i.e.:

$$K(x, y, t) = K_S(x, y) K_T(t), \tag{3}$$

where each part of the kernel is defined according to classical models of retinal processing [5, 24], namely DOG for the spatial part and a polynomial multiplied by a decaying exponential for the temporal part:

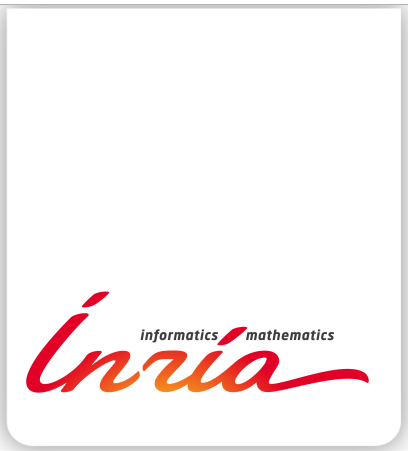
$$\begin{aligned} K_S(x, y) =& 16 \frac{1}{2\pi\sigma_c^2} \exp\left(-\frac{1}{2\sigma_c^2}\left((x - x_x)^2 + (y - x_y)^2\right)\right) \\ & - 8 \frac{1}{2\pi\sigma_s^2} \exp\left(-\frac{1}{2\sigma_s^2}\left((x - c_x)^2 + (y - c_y)^2\right)\right), \\ K_T(t) =& \left(-\frac{(0.7t)^7}{7!} + \frac{(0.7t)^5}{5!}\right) \exp\left(-0.7t\right), \end{aligned} \tag{4}$$

where parameters σ_c, σ_s define the spatial integration properties of the neuron (for center and surround) and (c_x, c_y) is the position of its center. These are the parameters that we vary to define the population.

Stage 2 gives the instantaneous spike rate $\lambda(t)$ by passing the output of the first stage by a non linearity:

$$\lambda(t) = f(L(t)), \text{ where } f(L) = \frac{1}{1 + \exp\left(-0.05L - 100\right)}.$$

Stage 3 converts the spike rate into a series of spikes using an inhomogeneous Poisson process. The time of the spikes is discretized into time bins of $1ms$.



RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399